

A non-extensive maximum entropy based regularization method for bad conditioned inverse problems

L. Rebollo-Neira and J. Fernandez-Rubio

Departament de Teoria del Senyal i Comunicacions,

Escola Tecnica Superior

d'Enginyers de Telecomunicació, Campus Nord, UPC,

Edifici D-4, c/. Gran Capita s/n. 08034, Barcelona, Spain

A. Plastino

Departamento de Física, Universidad Nacional de La Plata

C.C. 727, 1900 La Plata, Argentina.

Abstract

A regularization method based on the non-extensive maximum entropy principle is devised. Special emphasis is given to the $q = 1/2$ case. We show that, when the residual principle is considered as constraint, the $q = 1/2$ generalized distribution of Tsallis yields a regularized solution for bad-conditioned problems. The so devised regularized distribution is endowed with a component which corresponds to the well known regularized solution of Tikhonov.

1 Introduction

Consider a linear mapping \hat{A} from the space X of all distributions $\{p\}$ characterizing the statistics of a sample *into* the space Y of all measurable quantities. If one is given the distribution, the measurable quantities can be predicted. This is the so-called *direct* problem, a typical application being the computation of expectation values given a system's wave-function. Usually, a more difficult problem is posed by the inverse problem: given the measurable quantities, determine the underlying distribution $\{p\}$. In this paper we will concern ourselves with questions concerning the inverse problem, and will focus attention upon $p^{1/2}$ -distributions associated to the generalized non-extensive maximum entropy formalism of Tsallis (corresponding

to $q = 1/2$) [1, 2]. A physically relevant $q = 1/2$ instance is that related to the relaxation of two-dimensional Euler turbulence [3]. In the relaxation of Euler turbulence several identifiable stages can be identified: an initially hollow vorticity profile develops a linear diocotron instability which saturates with the creation of long-lived vortex patches. These patches move about for hundreds of diocotron periods, shedding filaments, and eventually mixing and inwardly transporting. This process gives rise to an axisymmetric metaequilibrium state, whose density decreases monotonically with radius, which then persists for tens of thousands of diocotron periods [3]. The shape and radial vorticity profile of the metaequilibrium state is an interesting and fundamental problem. As such, one could anticipate that it could be described by a variational principle. The most natural one applying within this context is Jaynes' maximum entropy principle, using Boltzmann-Gibbs' entropy. This yields, however, unsatisfactory results in this case [3]. Non-extensive entropy, with $q = 1/2$, allows for the use of Jaynes' variational principle and gives a convincing explanation of this phenomenon [3].

We are interested in the so-called stability issue of the inverse problem concerning the generalized non-extensive maximum entropy distribution corresponding to $q = 1/2$, which forces one to look also at the concomitant "direct" problem by assuming that it is *appropriately* represented by a linear mapping \hat{A} from the space X of all $p^{1/2}$ distributions characterizing the statistics of a sample, into the space Y of all corresponding measurable quantities.

In the present effort considerations will be restricted to the analysis of samples composed by a finite number of, say N , subsamples of known properties, so that the X -space will be identified with an N -dimensional Euclidean space $\mathbf{X}_N = \{p_j^{1/2} \geq 0; j = 1, \dots, N; \sum_{j=1}^N p_j = 1\}$. In practice the outcome of an experiment is a finite set of, say M , real numbers, so that the experimental measurements can be viewed as the components of an M -dimensional vector that we will call the *data vector* $|f\rangle \in \mathcal{R}^M$. The probability distribution can also be regarded as a

vector $|p^{\frac{1}{2}}\rangle$ [4]. Thus, the operator representing the direct problem admits a matrix representation of M rows and N columns.

In these terms the problem we wish to investigate consists of the following ingredients:

- When we apply the operator \hat{A} to all vectors $|p^{\frac{1}{2}}\rangle$ of \mathbf{X}_N we obtain a set of *images* that will be called the set of *exact* or *noise-free images* $|f\rangle$ verifying: $|f\rangle = \hat{A}|p^{\frac{1}{2}}\rangle$.
- On the other hand, in the concomitant inverse problem we are given an image $|f\rangle$ and are asked to find a $|p^{\frac{1}{2}}\rangle \in \mathbf{X}_N$ such that $|f\rangle = \hat{A}|p^{\frac{1}{2}}\rangle$.
- But, as a matter of fact, physical data are never known *exactly* but only with a certain degree of accuracy. In practice, rather than the *exact* image $|f\rangle$ the experimental setup gives rise to a vector $|f^o\rangle = |f\rangle + |\delta f\rangle$, where $|\delta f\rangle$ is a stochastic vector representing random noise.
- The impossibility of knowing the *exact* data may cause *serious* instabilities in the inverse problem solution, even if the operator \hat{A} does have an inverse. Indeed, the propagation of relative small errors from the data to the solution will often produce a meaningless solution if the so-called *condition number* $cond(\hat{A})$ is too large [5, 6, 7].
- The *condition number* is $cond(\hat{A}) = \mu_{max}/\mu_{min}$, where μ_{max} and μ_{min} are, respectively, the maximum and minimum singular values of the operator \hat{A} . A small variation $|\delta f\rangle$ of the data imposes bounds to the error of the solution according to [5, 6, 7]:

$$||\delta|p^{\frac{1}{2}}\rangle||_2/|||p^{\frac{1}{2}}\rangle||_2 \leq cond(\hat{A})||\delta f\rangle||_2/||f\rangle||_2,$$

where $||\cdot||_2$ indicates the 2-norm. As a consequence, when $cond(\hat{A})$ is too large, small fluctuations in the data can drastically affect the solution, whereby the problem is said to be *bad-conditioned* [5, 6].

The question is: how to achieve robustness (stability) of the solution against noise?

The answer: the solution needs to be *regularized*.

Consistently with the formulation of the direct problem, which is assumed to involve the $|p^{\frac{1}{2}}\rangle$ generalized distribution, we shall show in the present Communication that maximization of the $q = 1/2$ non-extensive entropy *renders a regularization method*, provided that a suitable constraint is imposed. *Regularization methods* have been the subject of much research during the last thirty years, [5] providing one with a good summary on the state of the art and a great deal of references.

In a recent publication [4] the generalized $|p^{\frac{1}{2}}\rangle$ non-extensive maximum entropy distribution was shown to yield a predictor of the data vector that minimizes the distance to the *real* (corrupted) data. In [4] linear-independence of the measurements is assumed. In the present effort we delve into the following points

- We shall prove that, provided that $Rank(\hat{A})$ is well defined, i.e, the spectrum of singular values of \hat{A} has a sharp cut-off, *the lack of linear-independence of the measurements does not affect the uniqueness of the $|p^{\frac{1}{2}}\rangle$ distribution.*
- Alas, when the spectrum of singular values has a fast decay rate, the $|p^{\frac{1}{2}}\rangle$ distribution minimizing the distance to the data may not exist when the data are perturbed by small errors. In order to obtain a stable (regularized) solution against small perturbations of the data we study the possibility of using as a constraint the *discrepancy or residual principle* [8, 9, 10] under the requirement

$$||\hat{A}|p^{\frac{1}{2}}\rangle - |f^o\rangle||_2 \leq |||\delta f\rangle||_2.$$

The paper is organized as follows: The inversion problem to be addressed is introduced in Section 2 via the description of the concomitant direct problem. In Section 3 the non-extensive maximum entropy approach is summarized and the uniqueness of the $|p^{\frac{1}{2}}\rangle$ distribution for the case of dependent measurements is discussed, while, in Section 4, the residual principle is shown to be an appropriate constraint that, together with the non-extensive maximum entropy principle, leads to a regularization method for badly conditioned inverse problems. For the $q = 1/2$ case, the regularized $|p^{\frac{1}{2}}\rangle$ distribution is shown to be the sum of two components, one of which corresponds to the well known regularized solution of Tikhonov [11]. The proposed regularization approach is illustrated in Section 5 by recourse to a numerical simulation concerning the determination of the statistical properties of a rare earth sample on the basis of magnetization measurements [12, 13].

2 The direct vs the inverse problem

As stated in the previous section, we shall focus attention on the stability aspect of the $|p^{\frac{1}{2}}\rangle$ distribution. For the sake of completeness we mention first the “direct” problem, corresponding to the inverse problem to be addressed, in terms of the general $|p^q\rangle$ distribution.

Let us assume that we wish to investigate a physical system S which is composed by N subsystems S_j ; $j = 1, \dots, N$ of known properties. In line with Tsallis’ proposal [1], let us further assume that the statistical description of such subsystems is given by the generalized weights $p_j^q \geq 0$; $j = 1, \dots, N$ such that $\sum_{j=1}^N p_j = 1$ [2]. Consider now that one interacts with the system S by means of a signal (probe) \mathbf{I} . If the physical laws ruling the interaction are well known then one could “predict” the outcome \mathbf{f} of the experiment, provided that the statistical distribution of subsystem is also known a priori. Thus, the above-mentioned “direct” problem

can be formulated in mathematical form as follows:

$$\hat{W}\mathbf{I} = \mathbf{f} \quad (1)$$

where the linear operator \hat{W} (associated to S) portrays the effect that the system produces upon the input signal so as to originate the response \mathbf{f} . As S is a composition of N subsystems S_j and the subsystems' statistics is assumed to be represented by the figures p_j^q ; $j = 1, \dots, N$ we can decompose the operator \hat{W} in the following fashion

$$\hat{W} = \sum_{j=1}^N p_j^q \hat{W}_j, \quad (2)$$

where the operator \hat{W}_j accounts for the (assumedly known) action of S_j upon the probe \mathbf{I} , i.e.,

$$\hat{W}_j \mathbf{I} = \boldsymbol{\alpha}_j \quad (3)$$

$\boldsymbol{\alpha}_j$ being the response evoked by S_j when impinged upon by the probe \mathbf{I} .

According to (1) and (2) and (3) we can write:

$$\hat{W}\mathbf{I} = \sum_{j=1}^N p_j^q \boldsymbol{\alpha}_j = \mathbf{f} \quad (4)$$

and it becomes clear that if we know both the figures p_j^q ; $j = 1, \dots, N$ and the physical laws governing the interaction, we can predict \mathbf{f} and the “direct” problem is solved.

On the other hand, the concomitant “inverse” problem concerns the “mirror” situation: knowing the response \mathbf{f} evoked by a system when impinged upon by a probe, determine the statistic distribution p_j^q ; $j = 1, \dots, N$ characterizing the sample (system) S . Since in this case \mathbf{f} is an experimental result, its available representation will be given as a finite set of, say M , numbers f_1, f_2, \dots, f_M and equation (4) will entail:

$$f_i = \sum_{j=1}^N p_j^q \alpha_{i,j}, \quad (5)$$

where $\alpha_{i,j}$ are the values one would obtain by performing the observation i upon the response α_j of system S_j . Certainly, the $\alpha_{i,j}$ values should also be derivable by recourse to physical considerations, as we are assuming that we deal with N subsystems S_j ; $j = 1, \dots, N$, of known properties. They can be regarded as matrix elements of an operator \hat{A} , i.e. $\alpha_{i,j} = \langle i | \hat{A} | j \rangle$; $i = 1, \dots, M$; $j = 1, \dots, N$ (the notation $|j\rangle$ is used to represent an element of the *standard basis* in Euclidean spaces). With this notation, and adopting the vectorial representation

$$|f\rangle = \sum_{i=1}^M f_i |i\rangle \quad (6)$$

the previous equation can be recast in the more compact form

$$|f\rangle = \hat{A} |p^q\rangle, \quad (7)$$

where $|p^q\rangle$ stands for the vectorial representation of p_j^q ; $j = 1, \dots, N$, i.e.:

$$|p^q\rangle = \sum_{j=1}^N p_j^q |j\rangle. \quad (8)$$

As already discussed, when one aims to solve the inverse problem, rather than the exact data $|f\rangle$, what is actually available is an observed vector $|f^o\rangle = |f\rangle + |\delta f\rangle$. How to take into account the uncertainty of the observed data so as to guarantee stability of the inverse problem-solutions will be the subject of Section 4.

3 The non-extensive maximum entropy approach

3.1 The generalized p^q distribution

The non-extensive maximum entropy approach for determining the generalized distribution p_j^q ; $j = 1, \dots, N$ entails maximizing the entropy S_q :

$$S_q = \frac{\sum_{j=1}^N p_j^q - \sum_{j=1}^N p_j}{1 - q} \quad (9)$$

subjected to the constraints of the problem. In the problem we are dealing with, if one wishes to reproduce the data $|f^o\rangle$, Tsallis' proposal leads to the maximization of (9) with constraints:

$$|f^o\rangle = \hat{A}|p^q\rangle \quad (10)$$

$$\sum_{j=1}^N p_j = 1. \quad (11)$$

This results in a distribution $|p^q\rangle$ of the form

$$p_j^q = z[1 - (1 - q)\langle j|\hat{A}^\dagger|\lambda\rangle]^{\frac{q}{1-q}} \quad ; j = 1, \dots, N, \quad (12)$$

where \hat{A}^\dagger stands for the adjoint of \hat{A} while z and the Lagrange Multiplier-vector $|\lambda\rangle \in \mathcal{R}^{\mathcal{M}}$ are determined so as to fulfill constraints (10) and (11).

In order to stress the role of the q -parameter we believe to be interesting to analyze the above optimization process by looking at it in a new, but equivalent, fashion: after the change of variables $p_j^q \rightarrow \tilde{p}_j$ one sees that, from a numerical viewpoint, the maximization of (9) with constraints (10) and (11) is tantamount to dealing with

$$\max \frac{\sum_{j=1}^N \tilde{p}_j}{1 - q} \quad (13)$$

subject to the constraints

$$|f^o\rangle = \hat{A}|\tilde{p}\rangle \quad (14)$$

$$\sum_{j=1}^N \tilde{p}_j^{\frac{1}{q}} = 1. \quad (15)$$

Indeed, by maximizing (13) with constraints (14) and (15) one obtains

$$\tilde{p}_j = z'[1 - (1 - q)\langle j|\hat{A}^\dagger|\lambda'\rangle]^{\frac{q}{1-q}} \equiv p_j^q \quad ; j = 1, \dots, N \quad (16)$$

with z' and $|\lambda'\rangle \in \mathcal{R}^{\mathcal{M}}$ determined by solving (14) and (15). We see then that the role of the q -parameter is completely specified by the constraint (15). If $q < 1$ such a constraint can be

expressed as $||\tilde{p}||_{\frac{1}{q}}^{\frac{1}{q}}$, where $||\tilde{p}||_{\frac{1}{q}}$ indicates the $1/q$ -norm. Thus, for $q < 1$, the selection of the q -parameter can be viewed as the process of choosing the norm one wishes to preserve while maximizing the 1-norm. The concomitant decision should, of course, be adopted by carefully paying attention to the physic of the problem at hand.

Notice that Boghosian chose an alternative path [3]: instead of constraint (11) he fixes: $\sum_{j=1}^N p_j^q$, and the corresponding maximum entropy scheme is then numerically equivalent to that of dealing with

$$\max -\frac{\sum_{j=1}^N \tilde{p}_j^{\frac{1}{q}}}{1-q} = \min \frac{\sum_{j=1}^N \tilde{p}_j^q}{1-q} \quad (17)$$

with the constraints

$$|f^o\rangle = \hat{A}|\tilde{p}\rangle \quad (18)$$

$$\sum_{j=1}^N \tilde{p}_j = 1, \quad (19)$$

whereby, for $q < 1$, maximizing the q -entropy is tantamount to minimizing $||\tilde{p}||_{\frac{1}{q}}^{\frac{1}{q}}$, while preserving the 1-norm. In this contribution we shall deal with constraint (11) and therefore, as already discussed, the parameter q is assigned the role of preserving the $1/q$ -norm. In particular, for $q = 1/2$ the Euclidean norm is seen to be preserved.

3.2 The $q = 1/2$ case

In a recent publication [4], some mathematical properties of the $q = 1/2$ distribution have been reported. Provided that the measurements $|f^o\rangle$ are independent, i.e. $Rank(\hat{A}) = M$, the $p_j^{\frac{1}{2}}$ distribution was shown to be given by

$$p_j^{\frac{1}{2}} = \sum_{n=1}^M \langle j|\psi_n\rangle \frac{1}{\sqrt{\mu_n}} \langle \tilde{\phi}_n|f^o\rangle + z_+ - z_+ \sum_{n=1}^M \sum_{k=1}^N \langle j|\psi_n\rangle \langle \psi_n|k\rangle \quad ; j = 1, \dots, N, \quad (20)$$

with $|\psi_n\rangle$ and $|\tilde{\phi}_n\rangle$; $n = 1, \dots, M$ satisfying the eigenvalue-equations

$$\hat{A}^\dagger A|\psi_n\rangle = \mu_n|\psi_n\rangle \quad ; \quad \mu_n > 0 \quad ; \quad n = 1, \dots, M, \quad (21)$$

$$\hat{A}\hat{A}^\dagger|\tilde{\phi}_n\rangle = \mu_n|\tilde{\phi}_n\rangle \quad ; \quad \mu_n > 0 \quad ; \quad n = 1, \dots, M, \quad (22)$$

The orthonormal vectors $|\psi_n\rangle$ and $|\tilde{\phi}_n\rangle$ are the so called *singular vectors* of \hat{A} and $\sqrt{\mu_n}$; $n = 1, \dots, M$ its non-zero *singular values*. The singular vectors are related each other through the mappings: $\hat{A}|\psi_n\rangle = \sqrt{\mu_n}|\tilde{\phi}_n\rangle$ and $\hat{A}^\dagger|\tilde{\phi}_n\rangle = \sqrt{\mu_n}|\psi_n\rangle$. The number z is calculated as

$$z = \left(\frac{1 - \sum_{n=1}^M \frac{|\langle \tilde{\phi}_n | f^o \rangle|^2}{\mu_n}}{N - \sum_{j=1}^N \sum_{k=1}^N \sum_{n=1}^M \langle j | \psi_n \rangle \langle \psi_n | k \rangle} \right)^{\frac{1}{2}}. \quad (23)$$

Although equation (20) was derived assuming $Rank(\hat{A}) = M$, it is straightforward to prove that it still holds true if $Rank(\hat{A}) = K < M$, with the corresponding summations running up to K instead of up to M . Indeed, if $Rank(A) = K < M$, in order to determine $p_j^{\frac{1}{2}}$; $j = 1, \dots, N$ we should obtain the Lagrange Multiplier-vector $|\lambda\rangle \in \mathcal{R}^{\mathcal{M}}$ by solving the equation

$$\hat{A}|p^{\frac{1}{2}}\rangle = \hat{A}|z\rangle - \frac{z}{2}\hat{A}\hat{A}^\dagger|\lambda\rangle = |f^o\rangle \quad (24)$$

$|z\rangle = \sum_{j=1}^N \langle j | z \rangle |j\rangle = \sum_{j=1}^N z |j\rangle$ being the vectorial representation of the real number z [4]. When $Rank(\hat{A}) = K < M$ the operator $\hat{A}\hat{A}^\dagger$ has no inverse and an infinite number of vectors $|\lambda\rangle \in \mathcal{R}^{\mathcal{M}}$ satisfying (24) exists. Nevertheless, the lack of uniqueness of the Lagrange multipliers does not affect the uniqueness of the distribution. This can be clearly appreciated if one writes the most general solution of equation (24), namely

$$|\lambda\rangle = \frac{2}{z}(\hat{A}\hat{A}^\dagger)^{-1}\hat{A}|z\rangle - \frac{2}{z}(\hat{A}\hat{A}^\dagger)^{-1}|f^o\rangle + |\lambda^\perp\rangle, \quad (25)$$

with $|\lambda^\perp\rangle$ an arbitrary vector in $Null(\hat{A}\hat{A}^\dagger)$ and $(\hat{A}\hat{A}^\dagger)^{-1}$ the pseudo-inverse of the operator $\hat{A}\hat{A}^\dagger$, which is expressed in terms of its eigenvectors as

$$(\hat{A}\hat{A}^\dagger)^{-1} = \sum_{n=1}^K |\tilde{\phi}_n\rangle \frac{1}{\mu_n} \langle \tilde{\phi}_n|. \quad (26)$$

Since

- $|\lambda^\perp\rangle \in Null(\hat{A}\hat{A}^\dagger)$ $\hat{A}^\dagger|\lambda^\perp\rangle = 0$, and

- $|p^{\frac{1}{2}}\rangle = |z\rangle - (z/2)\hat{A}^\dagger|\lambda\rangle$,

the arbitrary component $|\lambda^\perp\rangle$ has no effect whatsoever on $|p^{\frac{1}{2}}\rangle$. Thus, even if $\text{Rank}(\hat{A}) = K < M$ the $p_j^{\frac{1}{2}}(j = 1, \dots, N)$ distribution is of the form (20), with the summations running up to K instead of up to M .

From the above discussion we conclude that, at least in theory, the linear dependence of the measurements is overcome if the numerical Rank of operator \hat{A} can be estimated. However, difficulties of a fundamental nature raise their heads in many situations because, if the singular values spectrum does not have a clear cut off, the estimation of $\text{Rank}(\hat{A})$ is not an easy task. Furthermore, as it is well known, the existence of small singular values poses a bad-conditioned problem whereby small perturbations in the data cause a large dispersion in the solution [5, 6, 7]. One way to proceed, in order to restore stability to the solutions, might be that of applying the regularization method that prescribes the truncation of the singular values spectrum. We would like to stress that such a criterion is somewhat embodied in the $|p^{\frac{1}{2}}\rangle$ solution. Indeed, since the $|p^{\frac{1}{2}}\rangle$ -norm is fixed when, as a consequence of the data errors, the term $\sum_{n=1}^K \frac{|\langle \tilde{\phi}_n | f^o \rangle|^2}{\mu_n}$ (cf. eq (23)) becomes larger than unity there is no $|p^{\frac{1}{2}}\rangle$ solution. If one wishes for an approximate solution, the singular values-spectrum must be truncated so as to ensure a real value for the constant z . In the next section we shall present an alternative regularization method which is based on the *discrepancy or residual principle*.

4 The non-extensive maximum entropy based regularization method

We introduce here a constraint that allows for regularization of bad-conditioned inverse problems, namely, the residual principle [8, 9, 10]. This principle states that, rather than a predictor

that minimizes the distance to the data, we should look for a predictor, $\hat{A}|p^q\rangle$, such that:

$$||\hat{A}|p^q\rangle - |f^o\rangle||_2^2 \leq \delta^2, \quad (27)$$

where δ^2 is the square norm of the data error $|\delta f\rangle$. In order to determine the figures p_j^q ; $j = 1, \dots, N$, out of all the distributions satisfying condition (27) we will choose the one that, in addition to be endowed with the normalization property (11), maximizes the entropy (9). We incorporate constraint (27) into the variational process through the introduction of an additional variable $t \geq 0$ and write (27) as

$$||\hat{A}|p^q\rangle - |f^o\rangle||_2^2 + t = \delta^2 \quad t \geq 0. \quad (28)$$

The functional, \mathcal{L} , to be maximized is

$$\mathcal{L} = \frac{\sum_{j=1}^N p_j^q - \sum_{j=1}^N p_j}{1 - q} - \gamma_0 \sum_{j=1}^N p_j - \gamma(||\hat{A}|p^q\rangle - |f^o\rangle||_2^2 + t). \quad (29)$$

From the condition $\frac{\delta \mathcal{L}}{\delta t} = 0$ one immediately obtains that, either constraint (27) is irrelevant, or $t \equiv 0$. Assuming that the uniform distribution, which maximizes the unrestricted entropy, does not satisfy (27), the variable t must be zero and (27) becomes

$$||\hat{A}|p^q\rangle - |f^o\rangle||_2^2 = \delta^2. \quad (30)$$

From the condition $\frac{\delta \mathcal{L}}{\delta p} = 0$ it follows that

$$p_j^q = \gamma'_0 [1 - 2\gamma(1 - q)(\langle j|\hat{A}^\dagger A|p^q\rangle - \langle j|\hat{A}^\dagger|f^o\rangle)]^{\frac{q}{1-q}}, \quad (31)$$

where for the sake of convenience we have introduced $\gamma'_0 = [q/(1 + \gamma_0(1 - q))]^{\frac{q}{1-q}}$. In the next subsection we shall consider the case $q = 1/2$ and show that such a q -value considerably simplifies the computational task.

4.1 The $q = 1/2$ regularized solution

Adopting a $q = 1/2$ value for Tsallis' parameter, equation (31) is equivalent to

$$(\hat{I}\beta_1 + \hat{A}^\dagger A)|p^{\frac{1}{2}}\rangle = |\beta_2\rangle + \hat{A}^\dagger|f^o\rangle, \quad (32)$$

where \hat{I} is the identity operator in \mathcal{R}^N and we have set $\beta_1 = 1/(\gamma'_0\gamma)$ and $\beta_2 = 1/\gamma$, $|\beta_2\rangle = \sum_{j=1}^N \beta_2|j\rangle$ being the vectorial representation of the number β_2 . For $\beta_1 > 0$ the operator $(\hat{I}\beta_1 + \hat{A}^\dagger A)$ has an inverse that can be expressed as

$$(\hat{I}\beta_1 + \hat{A}^\dagger A)^{-1} = \sum_{n=1}^N |\psi_n\rangle \frac{1}{\beta_1 + \mu_n} \langle \psi_n|. \quad (33)$$

Hence, $|p^{\frac{1}{2}}\rangle$ can be cast in the fashion

$$|p^{\frac{1}{2}}\rangle = |c_{\beta_1}\rangle + \beta_2|\tilde{c}_{\beta_1}\rangle, \quad (34)$$

where

$$|c_{\beta_1}\rangle = \sum_{n=1}^N |\psi_n\rangle \frac{\sqrt{\mu_n}}{\beta_1 + \mu_n} \langle \tilde{\phi}_n|f^o\rangle, \quad (35)$$

and

$$|\tilde{c}_{\beta_1}\rangle = \sum_{n=1}^N \sum_{k=1}^N |\psi_n\rangle \frac{1}{\beta_1 + \mu_n} \langle \psi_n|k\rangle. \quad (36)$$

Notice that the $|c_{\beta_1}\rangle$ component of $|p^{\frac{1}{2}}\rangle$ corresponds to the regularized solution of Tikhonov [5, 6, 7, 11]. The remaining component, $|\tilde{c}_{\beta_1}\rangle$, appears as a consequence of the normalization constraint

$$\langle p^{\frac{1}{2}}|p^{\frac{1}{2}}\rangle = \sum_{j=1}^N p_j = 1. \quad (37)$$

Using this constraint to solve for β_2 one obtains two possible solutions:

$$\beta_{2\pm} = \frac{-\langle c_{\beta_1}|\tilde{c}_{\beta_1}\rangle \pm \sqrt{\langle c_{\beta_1}|\tilde{c}_{\beta_1}\rangle^2 - \langle \tilde{c}_{\beta_1}|\tilde{c}_{\beta_1}\rangle(\langle c_{\beta_1}|c_{\beta_1}\rangle - 1)}}{\langle \tilde{c}_{\beta_1}|\tilde{c}_{\beta_1}\rangle} \quad (38)$$

It is straightforward to prove that, for $\beta_1 > 0$, the β_{2-} solution is to be disregarded because it yields a lower entropy. Indeed, for β_2 satisfying (37) we have

$$S_{\frac{1}{2}} = 2 \sum_{j=1}^N \langle j | c_{\beta_1} \rangle + \beta_2 \langle j | \tilde{c}_{\beta_1} \rangle - 2. \quad (39)$$

Since $\beta_{2+} \geq \beta_{2-}$ the required proof is obtained by showing that $\sum_{j=1}^N \langle j | \tilde{c}_{\beta_1} \rangle \geq 0$. This immediately follows after writing $\sum_{j=1}^N \langle j | \tilde{c}_{\beta_1} \rangle$ in explicit fashion

$$\sum_{j=1}^N \langle j | \tilde{c}_{\beta_1} \rangle = \sum_{j=1}^N \sum_{k=1}^N \sum_{n=1}^N \langle j | \psi_n \rangle \frac{1}{\beta_1 + \mu_n} \langle \psi_n | k \rangle. \quad (40)$$

Thus, denoting $r_n = \sum_{j=1}^N \langle j | \psi_n \rangle$ it is clear that

$$\sum_{j=1}^N \langle j | \tilde{c}_{\beta_1} \rangle = \sum_{n=1}^N \frac{|r_n|^2}{\beta_1 + \mu_n} \geq 0. \quad (41)$$

The parameter β_1 , which characterizes the regularized solution (34), is to be fixed by solving equation (30).

5 Numerical Test

Consider that the sample to be analyzed is a mixture of $N = 11$ different rare earth elements. We shall deal with a simple paramagnetic model given in the literature [14, 15]. Our aim is that of determining the statistical composition of a paramagnetic sample on the basis of magnetization measurements.

We shall simulate the “direct” problem by assuming that the $|p^{\frac{1}{2}}\rangle$ distribution is of the binomial form

$$p_j^{\frac{1}{2}} = C \frac{11!}{(11-j)!j!} 0.4^j 0.6^{11-j} \quad j = 1, \dots, 11 \quad (42)$$

with C an appropriate constant accounting for the normalization condition (37).

We appeal now to a classical result: if, via a magnetic field H at the temperature T , one

interacts with a given ion j of Table I, one obtains the magnetization value [14, 15]:

$$\mathcal{M}_j(x) = g_j J_j \mu_B B_{J_j}(x) \quad (43)$$

where $x = H/T$, μ_B is the Bohr magneton, g_j is the spectral factor for the ion j

$$g_j = 1 + \frac{J_j(J_j + 1) + S_j(S_j + 1) - L_j(L_j + 1)}{2J_j(J_j + 1)} \quad (44)$$

and $B_{J_j}(x)$ is the Brillouin function

$$B_{J_j}(x) = \frac{2J_j + 1}{2J_j} \cotgh\left[\frac{2J_j + 1}{2J_j}x\right] - \frac{1}{2J_j} \cotgh\left[\frac{x}{2J_j}\right]. \quad (45)$$

For any given j we list the corresponding quantum number S_j, L_j and J_j in Table I.

Taking a series of $M = 40$ values the parameter $x_i = H_i/T$; ($i = 1, \dots, 40$) and using (42) we simulate the noise-free measurements as

$$f_i = \langle i | \hat{A} | p^{\frac{1}{2}} \rangle \quad ; i = 1, \dots, M, \quad (46)$$

where the matrix elements $\langle i | \hat{A} | j \rangle$ are the values $\mathcal{M}_j(x_i)$; $i = 1, \dots, 40$, ; $j = 1, \dots, 11$.

The inversion of equation (46) constitutes a typical example of a bad-conditioned problem. In fact, if the “exact data” f_i ; $i = 1, \dots, 40$ (calculated in double precision fashion) are given, using the inverse \hat{A}^{-1} one obtains the exact solution (42) with $Rank(\hat{A}) = 11$. However, as soon as we distort the exact data with small perturbations, the solution becomes chaotic. The continuous line in Fig 1 represents the exact solution given by (42) and also obtained through \hat{A}^{-1} from the exact data f_i ; $i = 1, \dots, M$ (see the continuous line of Fig 2 for a better resolution) . The broken and dotted lines of Fig 1 display the solution we obtain (for two different realizations) by randomly distorting the exact data within a 0.001% error.

In Fig 2 the continuous line represents the exact solution given in (42). The broken lines in the same figure represent the Tikhonov regularized solution (cf. eq (35)) from the same noisy data

as in the previous example. One appreciates now the fact that the two different realizations of the data yield similar results. The dotted lines of Fig 2 represent the $|p^{\frac{1}{2}}\rangle$ regularized solution. As it can be seen, both the Tikhonov regularization and the non-extensive maximum entropy regularization succeed in retrieving stability. In addition, the non-extensive maximum entropy regularized solution is shown to achieve a better agreement with the “true” solution.

6 Conclusions

Stability aspects of the generalized $q = 1/2$ Tsallis distribution, with regards to bad-conditioned problems, have here been addressed. It has been shown that, although redundant measurements do not affect the theoretical uniqueness of such a distribution in the noise free case, when dealing with problems of bad-conditioned nature small perturbation in the data require the regularization of the “idealistic” noise free solution.

On the basis of both the residual and the non-extensive maximum entropy principles, a regularization method has been presented in this Communication. The ensuing regularized $|p^{\frac{1}{2}}\rangle$ distribution is endowed with a component which corresponds to the well known regularized solution of Tikhonov. By recourse to a numerical test we were able to illustrate the fact that, in addition to guarantying stability against perturbations in the data, the $|p^{\frac{1}{2}}\rangle$ regularized solution may be better, “resolution-wise”, than the regularized solution of Tikhonov.

Acknowledgments

L. Rebollo-Neira is a member of CICIPBA (Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina). This work has been supported by CICYT of Spain (TIC95-1022-C05-1, TIC96-0500-C10-01) and CIRIT of Catalunya (1996 GR-00096). A. Plastino thanks CONICET for financial support.

Figure Captions

Table 1: For each rare-earth ion the pertinent quantum numbers are given.

Figure 1: The continuous line corresponds to the true distribution given in (35). The dotted and broken lines are the results (for two different realizations) obtained through \hat{A}^{-1} when the exact data are randomly distorted within a 0.001% error.

Figure 2: The continuous line corresponds to the true distribution given in (35). The broken lines display the regularized Tikhonov approximation for the same data as in Fig 1. The dotted lines are the corresponding non-extensive maximum entropy regularized results.

References

- [1] C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).
- [2] A. R. Plastino and A. Plastino, *Phys. Lett. A* **177**, 177 (1993).
- [3] B. M. R. Boghossian, *Phys. Rev. E* **53**, 4754 (1996).
- [4] L. Rebollo-Neira, A. Plastino, J. Fernandez-Rubio, *Physica A*, in press, (1998).
- [5] M. Bertero, *Advances in Electronics and Electron Physics*, ed P W Hawkes (New York Academic Presse), **75**, 1 (1989).
- [6] C. De Mol, *Inverse Problems in Optics*, Bellingham: SPIE, 345 1987.
- [7] C. W. Groetsch, *Inverse problems in the mathematical sciences*, Braunschweig-Wiesbaden: Vieweg, (1993)
- [8] D. L. Phillips, *J. Assoc. Comput. Mach.*, **9**, 84 (1962).
- [9] V. K. Ivanov *Comput. Math. Math. Phys.*, **6**, 6, 197 (1966)
- [10] V. A. Morozov, *Comput. Math. Math. Phys.*, **8**, 63 (1968).
- [11] A. N. Tikhonov, V. Y. Arsenin, *Solutions of Ill-Posed problems*, V. H. Winston & Sons, Washington, D.C., 1977.
- [12] A. Plastino, L. Rebollo-Neira, A. G. Alvarez, *Phys Rev. A* **40**, 1644 (1989).
- [13] L. Rebollo-Neira, A. Plastino, *Maximum Entropy and Bayesian Methods*, ed A. Mohammad-Djafari and G. Demoment, Kluwer Academic Publishers, Dordrecht, 67, (1993).

- [14] C. Kittel, *Introduction to Solid State Physics*, John Wiley and sons, Inc, New York (1959)
- [15] F. Reif, *Fundamentals of statistical and thermal physics*, McGraw-Hill:New York London, 1965